DOCUMENT RESUME

ED 325 059 HE 023 981

AUTHOR Johnson, William L.; Johnson, Annabel M.
TITLE Planning for University Faculty Assessment: Development of a Brief Summative-Evaluation Instrument.
PUB DATE 90
NOTE 89p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Accountability; College Faculty; *Evaluation Methods; *Evaluation Research; *Faculty Evaluation; Higher Education; *Summative Evaluation; Universities; Validated Programs

ABSTRACT

A description is provided of the process used in developing a brief summative-evaluation instrument that was created for use at a large university in the Southwestern United States in its attempt to be more responsive to the changing climate of assessment and accountability that universities are facing today. Beginning with a conceptual systemic model for addressing university faculty development, the discussion continues with the development of the instrument within the context of the conceptual model. The instrument was shown to be reliable and valid through the use of the Kuder-Richardson generalized reliability formula, coefficient alpha. Also included is a literature review for instrument development plus 19 attached appendices of data tables for the sampled participants from the areas of educational psychology, elementary education, educational administration, educational specializations, and secondary education. Contains 54 references. (GLR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED325059

PLANNING FOR UNIVERSITY FACULTY ASSESSMENT: DEVELOPMENT OF A BRIEF SUMMATIVE-EVALUATION INSTRUMENT

William L. Johnson, Ed.D.
Chairman, Education & Psychology
Ambassador College
Big Sandy, TX 75755

Annabel M. Johnson, Ph.D.
Ambassador College
Big Sandy, TX 75755

HE 023 981

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William Johnson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Abstract

This document describes the development of a brief summative–evaluation instrument at a large university in the Southwestern United States. Beginning with a conceptual systemic model for addressing university faculty development, this document discusses the development of the instrument within the context of the conceptual model. Also included is a literature review for instrument development plus attached appendices for the sampled participants from the areas of educational psychology, elementary education, educational administration, educational specializations and secondary education. The total number of participants for this study was 3003.

PLANNING FOR UNIVERSITY FACULTY ASSESSMENT: DEVELOPMENT OF A BRIEF SUMMATIVE-EVALUATION INSTRUMENT

Universities are addressing new issues today as they seek to become more responsive to a changing climate of assessment and accountability. In this document, we present a conceptual schema for addressing university development as well as a description of the development of a brief summative faculty-evaluation instrument. The instrument was developed for use in a large College of Education at a populous state university in the Southwestern United States.

The conceptual model presented grows out of an ecological approach to college developmental (i.e., a systems approach) and is well grounded in the research and management literature (Snyder & Anderson, 1986). The model functioned as a basis for our surveying the training desires of a college faculty within a context of changing expectations and new role demands.

Moving beyond the traditions of scientific management and human relations requires that deans and department heads consider the more dynamic concepts in temporary systems, contingency theory, matrix organization and organizational development.

An extensive literature review was conducted on effective schools, productive industries and social agencies to identify work patterns found in dynamic work cultures (Johnson & Snyder, 1986; Snyder, 1988; Snyder & Anderson, 1986). Given the scarcity of knowledge about excellence in colleges of education, the authors made certain assumptions. Viewing colleges as organizations, the authors postulated the same kinds of work patterns could invigorate faculty performance.

We found that school improvement planning, people and program development, and school assessment are important elements in productive organizations of all kinds. Consider the following organizational patterns.

Research Patterns Found in Productive Organizations

Organizational Planning Cluster

An analysis of various reports verifies that planning in productive organizations is collaborative in nature. Planning tasks include setting organizational development goals that relate to the organization's main mission of instruction (i.e., goal-tasks are assigned to ad hoc and permanent groups which plan collectively, holding individuals accountable for their role in the group's success) (Brookover & Lezotte, 1979; Edmonds, 1979; Glenn, 1981; Kanter, 1983; Lieberman & Miller, 1978; Peters & Waterman, 1982). An organizational plan becomes the focus for work, development and assessment.

Staff Development Cluster

Staff development programs that are linked to organizational improvement goals are the most effective. Staff development is an essential variable to improvement efforts and is planned by faculty and administrators together to address skills that are transferable to the work place (classroom) (Glenn, 1981; Lieberman & Miller, 1978; Miller & Wolf, 1978). On the job coaching and quality control systems enable professionals to develop new capacities and to align performance with goal targets. Further, when work groups – the building blocks of successful organizations – are provided with skill-building opportunities,

building blocks of successful organizations – are provided with skill-building opportunities, the capacity for shared inquiry and problem solving is enhanced. Collaborative forms of quality control are viewed as developmental and provide adjustment opportunities for the organization.

Program Development Cluster

In effective schools, educational leaders communicate instructional standards, coordinating programs and instruction. The faculty plans and carries out programs together, providing a climate of high achievement expectations for all students. The instructional program is characterized by adaptability and consistency in general, by clear and timely instructional cues, reinforcement, correctives and feedback, and by active student participation (Coulson, 1977; Edmonds, 1979; Klausemeier, 1982; Marcus, MacQueen, Duck, Lee & Wellisch, 1976).

Organizational Assessment Cluster

Managers of successful schools and other organizations develop organizational accountability models. Productive organizations have complex assessment systems that measure the success of goals. A complex educational assessment system, for example, includes measures of student achievement, work group productivity, staff performance and the outcomes of the organization's goals. Students in effective institutions are recognized for their accomplishments (Brookover & Lezotte, 1979; Edmonds, 1982). Assessment data in productive organizations provide a feedback loop for short-range planning and long-range

Organizational Culture Cluster

The best sellers on modern management now focus on making groups more productive and on empowering them to tackle problems and challenges consistently and persistently. The concept of corporate transformation (Kilmann, Saxton, & Serpa, 1985) has emerged to describe this current phenomenon in industry.

From the emerging work-culture literature (Kanter, 1983), we know that culture is so powerful that it can either stimulate or repress performance. That is, the strength of an organization's goals and symbol systems, the quality of group sharing and networking, the patterns of reward and recognition, and empowerment and entrepreneurial opportunities to shape the organization's future function as a context (work culture) that nurtures success. It seems that work culture is the social and psychological force that stimulates the direction and quality of work in an organization. The educational work place can be revitalized by stimulating and managing its culture.

The expansion of the knowledge base about organizational and human productivity over the past five years indicates that until educational leaders and support staff together assume responsibility for schooling, achievement patterns are likely to remain unchanged. The sociological-management challenge is one of empowering groups to address educational productivity.

Indeed, the literature on all kinds of productive organizations continues to affirm clearly and strongly that employee involvement in shaping the organization's direction is essential to the very survival of the organization. Opportunity, information, and resources are the vital materials that fuel organizational productivity.

Considering this context relative to faculty development and faculty assessment, we developed a summative–evaluation instrument for use in the college/university community. Indeed, any long–term educational operation should be in a constant state of examination for self–improvement. For this reason, colleges and universities should continually provide for their faculty evaluation. Accordingly there should be some system of feedback designed to assess and evaluate on the basis of collected data. Those responsible for making decisions must have access to data which permit them to determine faculty performance. Without such information, faculty effectiveness often cannot improve.

Recently a program evaluation committee in the College of Education at a large Southwestern State University (student enrollment of 25,000) developed a faculty–evaluation scale. Using the input of faculty and students, an assessment scale was designed. The final draft of the faculty and student survey instrument consisted of a set of fifteen teacher competencies thought to be important elements of effective teaching. The degree of emphasis was shown by a rating scale with the greater emphasis being shown by the higher rating.

Instrument Design and Methodology

The construction and field testing of an instrument involves the following steps: (a) development of a pool of instrument questions, (b) development of questionnaire forms, (c) selection of individuals to respond to the instrument, (d) actual administration of the instrument, and (e) writing computer programs for the reliability and validity assessments

of the instrument.

The questions were written using the standard domain-sampling model supported by Nunnally (1967) and Cronbach (1970). Initial questions addressing the teacher competency assessed were written, scrutinized, and studied. Next, the pool of items was expanded. Finally, the questions were rewritten and revised. These final revisions produced the pool of items contained in the instrument.

An instrument form employing a Likert format was developed for the pool of items. Lemon (1973) wrote that much of the experimental research pertaining to attitude change involves measuring attitudes by a simple, easily administered instrument, usually a rating-scale format, and then attributing changes or differences in the ratings to experimental manipulations. In this context, the Likert scale of summative ratings is the most widely used attitude measure in current research. Gronland (1976) stated that Likert's approach to attitude scaling was less time consuming in development than earlier methods such as Thurston's and Remmers'. Borg and Gall (1974) identified the Likert technique as the easiest method presently available for developing scales needed in attitude research.

McGinnies (1970) clarified the procedures involved in constructing a Likert scale. The first step involves the writing of a number of statements dealing with the topic under investigation. Respondents then indicate their reaction to each question on an interval scale. Severy (1974) added that it is assumed when a Likert scale is developed that all written questions reflect the same attitudinal dimensions. Advantages of the Likert scale are: (1) that construction is less time consuming and easier than are other scaling techniques, and (2) that fewer items are required to obtain higher correlation coefficients (Ahmann & Glock,

1975; Borg & Gall, 1974; Gronland, 1976; Lemon, 1973; Shaw & Wright, 1967).

Shaw and Wright (1967) wrote that items should be written in simple and clear language, should contain a single idea, and should be unambiguous. Borg and Gall (1974) noted that an instrument dealing with attitudes would generally be constructed as an attitude scale and should include at least ten items to obtain a valid picture of the attitude concerned. Tittle and Hill (1967) investigated the predictive efficiency of four frequently used attitude measurement techniques in terms of predicting indices of voter behavior. The Thurston, Guttman, Likert and semantic differential scales were compared. The Likert scale was superior to all other scales in predictive validity. Their evidence suggested that in cases where the Likert and Thurston scales were of equal length, the Likert scale exhibited higher reliability.

Jenkins and Taber (1977) found in an agree/disagree context that the number of response categories above five did not, in any situation, yield a significant increase in Likert discriminability. McKelvie (1978) proposed using five or six categories. He further suggested that there was no psychometric advantage in a large number of scale categories and, on the other hand, that discriminative power and validity may be reduced when fewer than five categories are used. Ramsey (1973) studied the effect of the number of categories in rating scales on the precision of scale values estimated by maximum likelihood techniques and concluded that using seven or more categories provides very nearly as much precision of estimate as a corresponding task requiring continuous judgment. Neumann and Neumann (1981) concluded from their research that the five-point scale appears to be the most convenient to use in attitudinal surveys. In addition to the fact that the literature suggestions

are frequent favoring a five- or six-point scale for the Likert format (Johnson, 1990), in our work we have found that the following six-category response choices recommended by Rotter (1972) seem to reflect relatively equidistant psychological order: 1=disagree strongly; 2=disagree; 3=tend to disagree; 4=tend to agree; 5=agree; and 6=agree strongly.

Instrument Scale Descriptors and Sampling Procedures

In constructing the response columns for the scale, the attempt was made to reflect equidistant order in the categories. In the continuum from "agree strongly" to "disagree strongly", the categories were ordered.

There is the question of a metric base for the variables. Stevens (1946) wrote that factor analysis requires the variables be measured on at least an interval level. This requirement is implied by the use of correlation or covariance matrices as the basic input to factor analysis. If it were argued that the scaling used was not absolutely of a metric base, would a researcher always avoid factor analysis? Kim and Mueller (1978) have written that such is not necessarily so. Many variables, such as measure of attitudes and opinions, do not have a clearly established metric base. Furthermore, it is generally assumed that many ordinal variables maybe given numeric values without distorting the underlying properties. The final answer to the question of applying multi-variate analysis to a nonmetric variable base really hinges on two considerations: (1) how well the arbitrarily assigned numbers reflect the underlying true distances, and (2) the amount of distortion introduced in the correlation procedures (which become the basic input to factor analysis) by the distortions of scaling. Fortunately, the correlation coefficients are fairly robust with respect to ordinal

distortions in measurement (Kim, 1975; Labovitz, 1967, 1970). Hence, as long as one can assume that the distortions introduced by assigning numeric values to ordinal categories are not very substantial, treating ordinal variables as if they were metric variables can be justified (Kim & Mueller, 1978). The scale descriptions were written to reflect as near a metric base as possible. It was felt the distortions were minimal and the application of factor analysis to the data would be consistent with factor analytic models used in instrument development.

The number of cases used in the instrument validation was 3003. The standard is to have at least five times as many subjects as try-out items (Lemke & Wiersma, 1976; Nunnally, 1970). The respondents populations were as follows:
(1) all students (n=3003); (2) educational psychology students (n=376); (3) elementary education students (n=1007); (4) educational administration students (n=208); (5) educational specializations students (n=950); and (6) secondary education students (n=422).

Final Scale Analysis of the Instrument

The collected pool of data was submitted to factor and subscale analyses (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Instrument validity was established by two methods. First content validity involves essentially the systematic examination of the scale content to determine whether it covers a representative sample of the domain to be measured (Anastasi, 1976; Ebel, 1956; Huddleston, 1956; Lennon, 1956). Representative selection of items and sensible test construction are cited by Nunnally (1967) as the standard means of ensuring content validity.

Next, construct validity was established. The construct validity of a test is the extent

to which the test may be said to measure a theoretical construct or trait (Anastasi, 1976). Essential to the notion of construct validity is the internal consistency of the items (Nunnally, 1967). Of particular relevance to construct validity is factor analysis. Lemke and Wiersma (1976) have defined factor analysis as the refined technique for analyzing the interrelationship of behavior data, simplifying the description of data by reducing the number of necessary variables or dimensions (Anastasi, 1976; Kim & Mueller, 1978).

Nunnally (1967) noted that some researchers refer to construct validity as "factorial validity." Furthermore, factor analysis is an attractive method for evaluating validity because it focuses on the reliable components of collected data (Gorsuch, 1983). Thompson (1989) noted that the "common variance" represented by indices of association tended to represent reliable variance, and since it is from these indices that factors are extracted, it follows that factors tend to be constructed from the "true score" components of variables.

Determining the number of factors to extract from the correlation matrix is a fundamental decision in any analysis (Thompson & Borrello, 1986). Many researchers follow the recommendations of Guttman (1954) and extract all factors with eigenvalues greater than one. Other researchers feel the Scree test is superior to other methods for determining the number of factors to extract (Cattell, 1965; Linn, 1968; Tucker, Koopman & Linn, 1969). These two methods have now been rigorously studied, and most of the leaders in factor analysis favor the Scree test.

We used principal components analysis (Stevens, 1986) followed by varimax rotation. Scale items were retained in the instrument only if they had a factor loading greater then or equal to 0.40. The rule often used in this context is to consider factor loading less than

0.30 as not substantial (Gorsuch, 1983; Kim and Mueller, 1978).

Evaluation of the reliability of the instrument was accomplished through use of the Kuder-Richardson generalized reliability formula, coefficient alpha (Cronbach, 1951; Ebrel, 1965; Novick & Lewis, 1967). This formula was appropriate since a Likert scaling format was employed in the instrument form.

The factor loadings for the instrument questions, as well as the Cronbach alpha reliability measures follow. Table 1 summarizes the psychometric data for the population sampled. Table 2 reports the reliability measures for each sample.

TABLE I
Summary Psychometric Data For The Scale

Area	Subjects	Instrument Questions	Factors	Factor Eigenvalue
All Respondents	3003	1-15	1	9.07
Educational Psychology	376	1-15	1	8.33
Elementary Education	1007	1-15	1	10.22
Educational Administration	208	1-4,6,9-15	1	7.48
Educational Specializations	950	1,4-15	1	8.55
Secondary Education	422	1-15	1	8.01

TABLE 2
Reliability Measures For The Scale

Area	Subjects	Instruments Questions	Cronbach's Alpha
All Respondents	3003	1-15	0.95731
Educational Psychology	376	1-15	0.94807
Elementary Education	1007	1-15	0.96920
Educational Administration	208	1-4,6,9-15	0.92488
Educational Specializations	950	1,4-15	0.95161
Secondary Education	422	1-15	0.94296

For the educational administration area, the data suggest that the scale is comprised of questions 1 through 4, 6 and 9 through 15. For the educational specializations area, the data suggest that the scale consist of questions 1 and 4 through 15.

In addition, the appendices to this document report the following information for each population surveyed: (1) descriptive statistics for the instrument variables (questions); (2) the correlation matrix for the instrument variables; and (3) the factor matrix using principal factoring (PA2) with interactions. See appendices A through R for this information. Finally the instrument variables are reported in Appendix S.

Summary and Conclusion

This article illustrated the process for the development of an instrument for faculty evaluation. The process linked the literature review with a psychometric analysis of the collected data.

The instrument was shown to be reliable and valid. Because of the reliability measures, the instrument can be used either for individual or group assessment. If it were desired that norms be established for each of the populations studied, such would involve the use of converting and equating procedures to generate a normal distribution with a specified mean and standard deviation. This process would allow the generation of percentile reference points for the interpretation of individual scores. A sample could be checked with respect to student classification, socioeconomic level, or other information relevant to the study. For this particular study, however, norms were not developed. Such could be a point-of-departure for another study which would encompass a separate body of psychometric literature and statistical analyses.

References

Ahmann, J.S., & Glock, M.D. (1975). **Evaluating pupil growth.** Boston: Allyn & Bacon.

Anastasi, A. (1976). **Psychological testing** (4th ed.) New York: Macmilliam.

Borg, W.R., & Gall, M. (1974). **Educational research.** New York: David McKay Company.

Brookover, W.B., & Lezotte, L.W. (1979). **Changes in school characteristics coincident with changes in student achievement.** East Lansing: Michigan State University, College of Education. (ERIC Document Reproduction Service No. ED 181 005).

Cattell, R.B. (1965). Factor analysis: an introduction to essentials (I) the purpose and underlying models, (II) the role of factor analysis in research. **Biometrics,** 21, 190-215, 405-435.

Coulson, J.E. (1977, July). **Overview of the national evaluation of the emergency school and aid act.** Santa Monica, CA: System Development Corporation (ED 154 951)

Cronbach, J.E. (1951). Coefficient alpha and the internal structure of tests. **Psychometrika,** 16, 3, 297-334.

Cronbach, L.J. (1970). **Essentials of psychological testing.** New York: MacMillan.

Ebel, R.L. (1956). Obtaining and reporting evidence on content validity. **Educational and Psychological Measurement,** 16, 3, 169-282.

Ebel, R.L. (1965). **Measuring educational achievement.** Englewood Cliffs, NJ: Prentice-Hall.

Edmonds, R.R. (1979). Effective schools for the urban poor. **Educational Leadership** 37, 15-23.

Edmonds, R.R. (December 1982). Programs for school improvement: An overview. **Educational Leadership,** 40, 4-11.

Glenn, B.C. (1981). **What works? An examination of effective schools for poor black children.** Cambridge: Harvard University, Center for Law and Education

Gorsuch, R.L. (1983). **Factor analysis.** Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gronland, N.E. (1976). **Measurement and evaluation in teaching.** New York: MacMillan.

Guttman, L. (1954). Some necessary conditions for common factor analysis. **Psychometrika,** 19, 149-161.

Huddelston, E.M. (1956). Test development on the basis of content validity. **Educational and Psychological Measurement,** 16, 3, 283-293.

Jenkins, G.D., & Taber, T.D. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. **Journal of Applied Psychology,** 62, 392-398.

Johnson, W.L. (1990). Factorial structure of the Charles F. Kettering Climate Profile. **Psychological Reports,** 67, 67-74.

Johnson, W.L., & Snyder, K.J. (1986). Instructional leadership effectiveness: A research analysis and strategy. **Educational and Psychological Research,** 6,1, 27-47.

Kanter, R.M. (1983). **The change masters: Innovation for productivity in the American corporation.** New York: Simon & Schuster

Kilmann, R., Saxton, M.J. & Serpa, R. (1985). **Gaining control of the corporate structure.** San Francisco: Jossey-Bass.

Kim, J.O. (1975). Mulitvariate analysis of ordinal variables. **American Journal of Sociology,** 81, 261-298.

Kim, J.O., & Mueller, C.W. (1978). **Factor Analysis.** Beverly Hills, CA.: Sage Publications.

Klausemeier, H.J. (1982, September). "A research strategy for educational improvement." **Educational Digest,** 48, 1, 19-22.

Labovitz, S. (1967). Some observations on measurement and statistics. **Social Forces,** 46, 151-160.

Labovitz, S. (1970). The assignments of numbers to rank order categories. **American Sociological Review,** 35, 515-524.

Lemke, E., & Wiersma, W. (1976). **Principles of psychological measurement.** Chicago, IL: Houghton-Mifflin.

Lemon, N. (1973). **Attitudes and their measurement.** London: B.T. Batsford.

Lennon, R.T. (1956). Assumptions underlying the use of content validity. **Educational and Psychological Measurement,** 16, 3, 294-304.

Lieberman, A., & Miller, L. (1978). Introduction to **Staff development: New demands, new realities, new perspectives.** New York: Teachers College Press.

Linn, R.L. (1968). A Monte Carlo approach to the number of factors problem.

Psychometrika, 33, 37-71.

Marcus, A.C., MacQueen, A.H., Duck, G.A., Lee, R.D. & Wellisch, J.B. (1976, April 14-16). **Administrative leadership in a sample of successful schools from the National Evaluation of the Emergency School.** System Development Corporation. Paper presented at the annual meeting of the AERA Convention, San Francisco, CA.

Miller, L., & Wolf, T. (1978). Staff development for school change: Theory and practice. In A. Lieberman & L. Miller (eds.), **Staff development: New demands, new realities, new perspectives** (pp. 144-160). New York: Teachers College Press.

McGinnies, E. (1970). **Social Behavior: A functional analysis.** Boston: Houghton Mifflin Company.

McKelvie, S.G. (1978). Graphic rating scales: how many categories. **British Journal of Psychology,** 69, 185-202.

Neumann L., & Neumann, Y. (1981). Comparison of six lengths of rating scales: students' attitudes toward instruction. **Psychological Reports,** 48, 399-404.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. (1975). **Statistical package for the social sciences.** New York, NY: McGraw-Hill.

Novick, M.R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. **Psychometrika,** 32, 1, 1-13.

Nunnally, J.C. (1967). **Psychometric theory.** New York, NY: McGraw-Hill.

Nunnally, J.C. (1970). **Introduction to Psychological Measurement.** New York, NY: McGraw-Hill.

Peters, T.J., & Waterman, R.H. (1982). **In Search of Excellence: Lessons from America's Best Run Companies.** New York: Harper & Row.

Ramsey, J.O. (1973). The effect of number of categories in rating scales on precision of estimation of scale values. **Psychometrika,** 38, 513-532.

Rotter, G.S. (1972). Attitude points of agreement and disagreement. **Journal of Social Psychology,** 86, 211-218.

Severy, L.J. (1974). **Procedure and issues in the measurement of attitudes.** Washington, D.C.: National Institute of Education. (ERIC Document Reproduction Service No. ED 099 426).

Shaw, M.E., & Wright, J.M. (1967). **Scales for the measurement of attitudes.** New York: McGraw-Hill.

Snyder, K.J. & Anderson, R.H. (1986). **Managing productive schools: Toward an ecology.** San Diego: Harcourt Brace & Jovanovich, Inc.

Snyder, K.J. (1988). **Managing productive schools.** San Diego: Harcourt Brace & Jovanovich, Inc.

Stevens, J. (1986). **Applied Multivariate Statistics for the Social Sciences.** Hillsdale, N.J.: Lawrence Erlbaum & Associates, Inc.

Stevens, S.S. (1946). On the theory of scales measurement. **Science,** 103, 677-680.

Thompson, B. (1989). Meta-analysis of factor structures: a case study example with Bem's androgyny measure. **Journal of Experimental Education,** 57, 187-197.

Thompson, B., & Borrello, G.M. (1986). Second-order factor structure of the MBTI: a construct validity assessment. **Measurement and Evaluation in Counseling and Development,** 18, 148-153.

Tittle, C.R., & Hill, R.E.J. (1967). Attitude measurement and prediction of behavior: An evaluation of conditions and measurement techniques. **Sociometry,** 30, 2, 199-213.

Tucker, L.R., Koopman, R.F., & Linn, R.L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. **Psychometrika,** 34, 421-459.

APPENDIX A
Descriptive Statistics for the Instrument
Variables (Questions)
All Respondents (n=3003)

Variables	M	Mdn	Mo	SE	SD
Q1	3.419	3.610	4.000	0.014	0.780
Q2	3.506	3.617	4.000	0.012	0.661
Q3	3.502	3.648	4.000	0.013	0.695
Q4	3.416	3.598	4.000	0.014	0.794
Q5	3.526	3.687	4.000	0.013	0.699
Q6	3.538	3.718	4.000	0.013	0.725
Q7	3.570	3.718	4.000	0.012	0.662
Q8	3.717	3.847	4.000	0.011	0.589
Q9	3.483	3.648	4.000	0.013	0.730
Q10	3.475	3.636	4.000	0.013	0.726
Q11	3.375	3.541	4.000	0.014	0.787
Q12	3.360	3.526	4.000	0.014	0.793
Q13	3.405	3.619	4.000	0.015	0.816
Q14	3.505	3.690	4.000	0.014	0.741
Q15	3.444	3.690	4.000	0.016	0.851

APPENDIX B

Correlation Matrix for the Instrument Variables (Questions)
All Respondents (n=3003)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.4740	1.0000													
Q3	0.5256	0.4849	1.0000												
Q4	0.6230	0.4698	0.5438	1.0000											
Q5	0.6301	0.5195	0.6345	0.6391	1.0000										
Q6	0.6274	0.4950	0.6288	0.5878	0.7291	1.0000									
Q7	0.5713	0.4834	0.6825	0.5441	0.6908	0.6872	1.0000								
Q8	0.5618	0.4857	0.5450	0.5229	0.6184	0.6307	0.6059	1.0000							
Q9	0.6102	0.4658	0.5391	0.5818	0.6171	0.6324	0.5819	0.6587	1.00000						
Q10	0.6314	0.4850	0.5771	0.5758	0.6367	0.7100	0.6154	0.6367	0.66069	1.00000					
Q11	0.6042	0.5149	0.5324	0.6411	0.6317	0.6064	0.5872	0.5728	0.59338	0.62355	1.0000				
Q12	0.5936	0.4371	0.5355	0.6155	0.6166	0.5973	0.5737	0.5302	0.56577	0.58610	0.58784	1.0000			
Q13	0.6321	0.4542	0.5356	0.5996	0.6445	0.6706	0.5777	0.6258	0.68260	0.66729	0.60672	0.65482	1.0000		
Q14	0.6072	0.4470	0.5361	0.5809	0.6511	0.6748	0.5966	0.6324	0.69070	0.66747	0.58197	0.61715	0.81044	1.00000	
Q15	0.6642	0.4650	0.5689	0.6130	0.6758	0.7099	0.6191	0.6373	0.66280	0.69451	0.62555	0.62829	0.77446	0.75430	1.0000

APPENDIX C

Factor Matrix Using Principal Factoring (PA2) with Iterations
All Respondents (n=3003)

Variables	Factor 1*
Q1	0.769
Q2	0.604
Q3	0.719
Q4	0.746
Q5	0.825
Q6	0.833
Q7	0.774
Q8	0.760
Q9	0.788
Q10	0.810
Q11	0.762
Q12	0.748
Q13	0.830
Q14	0.821
Q15	0.845

*Eigenvalue of 9.07

APPENDIX D
Means and Standard Deviations for the Instrument
Variables (Questions)
Educational Psychology Area (n=376)

Variables	M	SD
Q1	3.5426	0.6915
Q2	2.5319	0.6274
Q3	3.5293	0.6999
Q4	3.4681	0.7859
Q5	3.5904	0.6510
Q6	3.6223	0.6661
Q7	3.6250	0.6494
Q8	3.7846	0.5048
Q9	3.5106	0.6927
Q10	3.5080	0.6850
Q11	3.4255	0.7191
Q12	3.4840	0.7301
Q13	3.5133	0.7999
Q14	3.5559	0.7133
Q15	3.5372	0.7720

APPENDIX E

Correlation Matrix for the Instrument Variables (Questions)
for the Educational Foundations Area (n=376)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.4947	1.0000													
Q3	0.4685	0.4077	1.0000												
Q4	0.5913	0.4239	0.4695	1.0000											
Q5	0.5897	0.5021	0.5999	0.5842	1.0000										
Q6	0.5155	0.4692	0.5386	0.5118	0.6816	1.0000									
Q7	0.4780	0.4385	0.6843	0.4389	0.6387	0.5718	1.0000								
Q8	0.4961	0.4385	0.4670	0.4363	0.5747	0.6060	0.4688	1.0000							
Q9	0.5835	0.3980	0.4476	0.5835	0.5063	0.5346	0.4920	0.5594	1.0000						
Q10	0.5199	0.4492	0.5001	0.4635	0.5395	0.6319	0.4893	0.5718	0.5420	1.0000					
Q11	0.5480	0.4781	0.4255	0.5572	0.5157	0.4922	0.5139	0.4515	0.4994	0.5127	1.0000				
Q12	0.5823	0.5075	0.5253	0.5335	0.6650	0.5962	0.5807	0.5079	0.5433	0.5360	0.5615	1.0000			
Q13	0.5895	0.4375	0.5185	0.5712	0.5891	0.6501	0.5358	0.6246	0.5892	0.6178	0.5001	0.6738	1.0000		
Q14	0.5708	0.3862	0.5308	0.5097	0.5891	0.6281	0.5778	0.5778	0.5788	0.6102	0.5461	0.6493	0.7791	1.0000	
Q15	0.6214	0.4655	0.5877	0.5865	0.6406	0.7016	0.5997	0.6125	0.5976	0.6625	0.5526	0.6586	0.8218	0.7879	1.0000

APPENDIX F

Factor Matrix using Prinicpal Factoring (PA2) with Iterations
for the Educational Psychology Area (n=376)

Variables	Factor 1*
Q1	0.733
Q2	0.595
Q3	0.686
Q4	0.694
Q5	0.792
Q6	0.786
Q7	0.719
Q8	0.710
Q9	0.714
Q10	0.735
Q11	0.681
Q12	0.784
Q13	0.829
Q14	0.811
Q15	0.870

*Eigenvalue of 8.23

APPENDIX G

Means and Standard Deviations for the Instrument Variables (Questions)
Elementary Education Area (n=1007)

Variables	M	SD
Q1	3.3118	0.8643
Q2	3.4300	0.7372
Q3	3.4012	0.7748
Q4	3.3128	0.8509
Q5	3.4320	0.7920
Q6	3.4141	0.8247
Q7	3.4677	0.7531
Q8	3.6058	0.7053
Q9	3.3774	0.8091
Q10	3.4022	0.7989
Q11	3.2721	0.8745
Q12	3.2175	0.8503
Q13	3.2890	0.8833
Q14	3.4320	0.7995
Q15	3.3178	0.9386

APPENDIX H

Correlation Matrix for the Instrument Variables (Questions) Elementary Education Area (n=1007)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.5964	1.0000													
Q3	0.6032	0.5834	1.0000												
Q4	0.7383	0.6029	0.6356	1.0000											
Q5	0.7214	0.6535	0.7005	0.7063	1.0000										
Q6	0.7327	0.6288	0.6777	0.6607	0.7849	1.0000									
Q7	0.6589	0.6166	0.7256	0.6400	0.7523	0.7632	1.0000								
Q8	0.6622	0.6378	0.6371	0.6279	0.7268	0.7354	0.7122	1.000							
Q9	0.6822	0.5908	0.6144	0.6599	0.7039	0.7144	0.6642	0.7590	1.0000						
Q10	0.7315	0.6174	0.6672	0.6920	0.7415	0.7818	0.7310	0.7491	0.7291	1.0000					
Q11	0.6891	0.6246	0.6411	0.7203	0.7213	0.7008	0.6939	0.6558	0.6666	0.7167	1.0000				
Q12	0.6670	0.5340	0.5795	0.6820	0.6573	0.6141	0.6046	0.5856	0.6087	0.6671	0.6408	1.0000			
Q13	0.6910	0.5828	0.5915	0.6810	0.7051	0.7211	0.6467	0.6840	0.7290	0.7211	0.6804	0.6692	1.0000		
Q14	0.6772	0.5918	0.6040	0.6675	0.7096	0.7218	0.6627	0.7270	0.7419	0.7252	0.6349	0.6307	0.8350	1.0000	
Q15	0.7398	0.5722	0.6268	0.7005	0.7324	0.7607	0.6823	0.7224	0.6900	0.7400	0.7191	0.6468	0.8075	0.7613	1.0000

APPENDIX I

Factor matrix using Principal Factoring (PA2) with Iterations for the Elementary Education Area (n=1007)

Variables	Factor 1*
Q1	0.83142
Q2	0.72402
Q3	0.76571
Q4	0.81388
Q5	0.87073
Q6	0.87025
Q7	0.82757
Q8	0.83465
Q9	0.82848
Q10	0.87085
Q11	0.82267
Q12	0.75682
Q13	0.84672
Q14	0.84167
Q15	0.86164

*Eigenvalue of 10.22

APPENDIX J

Means and Standard Deviations for the Instrument Variables (Questions)
Educational Administration Area (n=208)

Varibles	M	SD
Q1	3.5865	0.6311
Q2	3.6635	0.5665
Q3	3.5913	0.6455
Q4	3.4567	0.8094
Q5	3.6058	0.6046
Q6	3.6346	0.6454
Q7	3.6538	0.5777
Q8	3.8029	0.3988
Q9	3.6106	0.6647
Q10	3.5481	0.6502
Q11	3.4038	0.7226
Q12	3.4615	0.7540
Q13	3.5481	0.6934
Q14	3.6058	0.6281
Q15	3.6587	0.6622

APPENDIX K

Correlation Matrix for the Instrument Variables (Questions)
Educational Administration Area (n=208)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3655	1.0000													
Q3	0.4488	0.4411	1.0000												
Q4	0.2958	0.3262	0.4144	1.0000											
Q5	0.2924	0.1890	0.5135	0.4091	1.0000										
Q6	0.4575	0.3887	0.7297	0.4597	0.5451	1.0000									
Q7	0.3210	0.2032	0.6551	0.2467	0.5755	0.5790	1.0000								
Q8	0.4424	0.2181	0.4549	0.3102	0.5376	0.5259	0.5621	1.0004							
Q9	0.4089	0.3430	0.4717	0.4579	0.5056	0.5563	0.3266	0.4744	1.0000						
Q10	0.5078	0.4375	0.6397	0.4491	0.4662	0.6866	0.4946	0.5118	0.5297	1.0000					
Q11	0.4526	0.2982	0.3969	0.5008	0.4988	0.4422	0.3943	0.4787	0.5503	0.5137	1.0000				
Q12	0.3724	0.3201	0.5878	0.5157	0.4963	0.6459	0.4572	0.4325	0.5241	0.6146	0.4808	1.0000			
Q13	0.4099	0.3734	0.6323	0.4556	0.4717	0.6871	0.4880	0.5323	0.6644	0.6806	0.5010	0.6595	1.0000		
Q14	0.5008	0.3720	0.5776	0.4318	0.4410	0.7154	0.4075	0.4982	0.6718	0.6616	0.4589	0.6308	0.7868	1.0000	
Q15	0.5507	0.4134	0.5648	0.3373	0.3983	0.5884	0.3589	0.3659	0.5416	0.6161	0.4409	0.5202	0.6093	0.6389	1.000

APPENDIX L

Factor Matrix Using Principal Factoring (PA2) with Iterations for the Educational Administration Area (n=208)*

Variables	Factor 1
Q1	0.587
Q2	0.472
Q3	0.777
Q4	0.565
Q5	0.651
Q6	0.834
Q7	0.658
Q8	0.649
Q9	0.723
Q10	0.807
Q11	0.642
Q12	0.749
Q13	0.832
Q14	0.818
Q15	0.717

*Eigenvalue for Factor 1 is 7.48334.

APPENDIX M

Means and Standard Deviations for the Instrument
Variables (Questions)
Educational Specializations Area (n=950)

Variables	M	SD
Q1	3.4684	0.7462
Q2	3.5642	0.5946
Q3	3.5505	0.6047
Q4	3.4937	0.7254
Q5	3.5568	0.6512
Q6	3.5684	0.6739
Q7	3.6116	0.5894
Q8	3.7747	0.5232
Q9	3.5589	0.6779
Q10	3.4905	0.7166
Q11	3.4421	0.7246
Q12	3.4189	0.7649
Q13	3.4674	0.7588
Q14	3.5453	0.7269
Q15	3.4705	0.8343

APPENDIX N

Correlation Matrix for the Instrument Variables (Questions) Educational Specializations Area (n=950)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3441	1.0000													
Q3	0.4950	0.4158	1.0000												
Q4	0.6370	0.3478	0.5087	1.0000											
Q5	0.6627	0.4314	0.5961	0.6487	1.0000										
Q6	0.5993	0.3558	0.5680	0.5785	0.7018	1.0000									
Q7	0.5602	0.3944	0.6390	0.5426	0.6547	0.6386	1.0000								
Q8	0.5053	0.3242	0.4323	0.4682	0.5077	0.4830	0.5155	1.0000							
Q9	0.6149	0.3042	0.4849	0.5353	0.5999	0.5723	0.5623	0.5989	1.0000						
Q10	0.6279	0.3265	0.4825	0.5613	0.6063	0.6787	0.5513	0.5395	0.6583	1.0000					
Q11	0.5909	0.4378	0.5020	0.6207	0.5987	0.5616	0.5481	0.5242	0.5518	0.6005	1.0000				
Q12	0.6083	0.3114	0.4849	0.6257	0.6058	0.5964	0.5997	0.4809	0.5375	0.5493	0.5913	1.0000			
Q13	0.6698	0.3047	0.4467	0.5681	0.6584	0.6152	0.5312	0.5653	0.6612	0.6688	0.5839	0.6226	1.0000		
Q14	0.6183	0.2943	0.4574	0.5660	0.6846	0.6507	0.5858	0.5698	0.6659	0.6632	0.5600	0.6101	0.8058	1.0000	
Q15	0.6711	0.3224	0.4968	0.5976	0.6634	0.6650	0.5884	0.5858	0.6652	0.7096	0.5879	0.6186	0.7540	0.7630	1.0000

APPENDIX O

Factor Matrix Using Principal Factoring (PA2) With Iterations* for the Educational Specializations Area (n=950)*

Variables	Factor 1
Q1	0.78135
Q2	0.45678
Q3	0.66740
Q4	0.74328
Q5	0.82239
Q6	0.79184
Q7	0.75526
Q8	0.67143
Q9	0.76848
Q10	0.79062
Q11	0.74429
Q12	0.74838
Q13	0.82749
Q14	0.82983
Q15	0.84443

*Eigenvalue for Factor 1 is 8.55774.

APPENDIX P

Means and Standard Deviations for the Instrument Variables (Questions)
Secondary Education Area (n=422)

Variables	M	SD
Q1	3.3720	0.7591
Q2	3.4431	0.6649
Q3	3.5592	0.6858
Q4	3.4597	0.7209
Q5	3.5853	0.6362
Q6	3.6209	0.6453
Q7	3.6232	0.6069
Q8	3.7536	0.5435
Q9	3.4763	0.6948
Q10	3.5308	0.6300
Q11	3.4384	0.7485
Q12	3.3815	0.7606
Q13	3.3697	0.8276
Q14	3.4858	0.7053
Q15	3.4882	0.7970

APPENDIX Q

Correlation Matrix for the Instrument Variables (Questions)
Secondary Education Area (n=422)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3549	1.0000													
Q3	0.4069	0.3980	1.0000												
Q4	0.4636	0.4758	0.5164	1.0000											
Q5	0.5070	0.4241	0.5762	0.5978	1.0000										
Q6	0.4874	0.4090	0.5982	0.5593	0.6979	1.0000									
Q7	0.4544	0.3499	0.6215	0.5487	0.6462	0.6532	1.0000								
Q8	0.4069	0.4014	0.5299	0.4534	0.5005	0.5930	0.5314	1.0000							
Q9	0.4603	0.4777	0.5013	0.5291	0.5015	0.5996	0.5279	0.5694	1.0000						
Q10	0.4602	0.4634	0.5262	0.4812	0.4971	0.6130	0.5366	0.6771	0.6093	1.0000					
Q11	0.5023	0.4725	0.4420	0.5719	0.5821	0.5956	0.4898	0.5230	0.5566	0.5631	1.0000				
Q12	0.4200	0.3787	0.4642	0.5846	0.5043	0.4840	0.4407	0.4692	0.5137	0.4240	0.5065	1.0000			
Q13	0.4687	0.3836	0.4928	0.5585	0.5309	0.6322	0.5143	0.5673	0.6637	0.5430	0.5928	0.6357	1.0000		
Q14	0.4291	0.3908	0.4829	0.5407	0.5293	0.6143	0.5007	0.5405	0.6900	0.5355	0.5674	0.5436	0.7863	1.0000	
Q15	0.4607	0.4469	0.5075	0.5758	0.6296	0.6793	0.5824	0.5251	0.6557	0.5423	0.5681	0.5853	0.7304	0.7222	1.0000

APPENDIX R

Factor Matrix Using Prinicpal Factoring (PA2) With Iterations for the Secondary Education Area (n=422)*

Variables	Factor 1
Q1	0.609
Q2	0.561
Q3	0.690
Q4	0.729
Q5	0.759
Q6	0.815
Q7	0.727
Q8	0.705
Q9	0.775
Q10	0.721
Q11	0.738
Q12	0.680
Q13	0.805
Q14	0.783
Q15	0.816

*Eigenvalue for Factor 1 is 8.01242.

APPENDIX S

Questions on the Instrument

1. Has made his/her expectations clear.
2. Has maintained adequate office hours for consultation.
3. Has been tolerant of differing viewpoints regarding the subject matter.
4. Has utilized appropriate grading practices.
5. Has been responsive to student needs.
6. Has related well to students.
7. Has been open to student input (suggestions, comments, recommendations, etc.).
8. Has appeared knowledgeable in his/her field.
9. Has presented material relevant to the course description in the catalog.
10. Has communicated (speech, mannerisms, delivery, etc.) effectively.
11. Has provided appropriate feedback on student work.
12. Has given realistic class assignments.
13. Has made the course worthwhile (I have learned enough to make it worth my time and money).
14. Has presented the course in such a manner that I see present or future applications of the materials, ideas, skills, etc.
15. Has been of such caliber that I would recommend him/her to other students.

PLANNING FOR UNIVERSITY FACULTY ASSESSMENT: DEVELOPMENT OF A BRIEF SUMMATIVE–EVALUATION INSTRUMENT

William L. Johnson, Ed.D.
Chairman, Education & Psychology
Ambassador College
Big Sandy, TX 75755

Annabel M. Johnson, Ph.D.
Ambassador College
Big Sandy, TX 75755

Abstract

This document describes the development of a brief summative–evaluation instrument at a large university in the Southwestern United States. Beginning with a conceptual systemic model for addressing university faculty development, this document discusses the development of the instrument within the context of the conceptual model. Also included is a literature review for instrument development plus attached appendices for the sampled participants from the areas of educational psychology, elementary education, educational administration, educational specializations and secondary education. The total number of participants for this study was 3003.

PLANNING FOR UNIVERSITY FACULTY ASSESSMENT: DEVELOPMENT OF A BRIEF SUMMATIVE-EVALUATION INSTRUMENT

Universities are addressing new issues today as they seek to become more responsive to a changing climate of assessment and accountability. In this document, we present a conceptual schema for addressing university development as well as a description of the development of a brief summative faculty-evaluation instrument. The instrument was developed for use in a large College of Education at a populous state university in the Southwestern United States.

The conceptual model presented grows out of an ecological approach to college developmental (i.e., a systems approach) and is well grounded in the research and management literature (Snyder & Anderson, 1986). The model functioned as a basis for our surveying the training desires of a college faculty within a context of changing expectations and new role demands.

Moving beyond the traditions of scientific management and human relations requires that deans and department heads consider the more dynamic concepts in temporary systems, contingency theory, matrix organization and organizational development.

An extensive literature review was conducted on effective schools, productive industries and social agencies to identify work patterns found in dynamic work cultures (Johnson & Snyder, 1986; Snyder, 1988; Snyder & Anderson, 1986). Given the scarcity of knowledge about excellence in colleges of education, the authors made certain assumptions. Viewing colleges as organizations, the authors postulated the same kinds of work patterns could invigorate faculty performance.

We found that school improvement planning, people and program development, and school assessment are important elements in productive organizations of all kinds. Consider the following organizational patterns.

Research Patterns Found in Productive Organizations

Organizational Planning Cluster

An analysis of various reports verifies that planning in productive organizations is collaborative in nature. Planning tasks include setting organizational development goals that relate to the organization's main mission of instruction (i.e., goal-tasks are assigned to ad hoc and permanent groups which plan collectively, holding individuals accountable for their role in the group's success) (Brookover & Lezotte, 1979; Edmonds, 1979; Glenn, 1981; Kanter, 1983; Lieberman & Miller, 1978; Peters & Waterman, 1982). An organizational plan becomes the focus for work, development and assessment.

Staff Development Cluster

Staff development programs that are linked to organizational improvement goals are the most effective. Staff development is an essential variable to improvement efforts and is planned by faculty and administrators together to address skills that are transferable to the work place (classroom) (Glenn, 1981; Lieberman & Miller, 1978; Miller & Wolf, 1978). On the job coaching and quality control systems enable professionals to develop new capacities and to align performance with goal targets. Further, when work groups – the building blocks of successful organizations – are provided with skill-building opportunities,

building blocks of successful organizations – are provided with skill-building opportunities, the capacity for shared inquiry and problem solving is enhanced. Collaborative forms of quality control are viewed as developmental and provide adjustment opportunities for the organization.

Program Development Cluster

In effective schools, educational leaders communicate instructional standards, coordinating programs and instruction. The faculty plans and carries out programs together, providing a climate of high achievement expectations for all students. The instructional program is characterized by adaptability and consistency in general, by clear and timely instructional cues, reinforcement, correctives and feedback, and by active student participation (Coulson, 1977; Edmonds, 1979; Klausemeier, 1982; Marcus, MacQueen, Duck, Lee & Wellisch, 1976).

Organizational Assessment Cluster

Managers of successful schools and other organizations develop organizational accountability models. Productive organizations have complex assessment systems that measure the success of goals. A complex educational assessment system, for example, includes measures of student achievement, work group productivity, staff performance and the outcomes of the organization's goals. Students in effective institutions are recognized for their accomplishments (Brookover & Lezotte, 1979; Edmonds, 1982). Assessment data in productive organizations provide a feedback loop for short-range planning and long-range

Organizational Culture Cluster

The best sellers on modern management now focus on making groups more productive and on empowering them to tackle problems and challenges consistently and persistently. The concept of corporate transformation (Kilmann, Saxton, & Serpa, 1985) has emerged to describe this current phenomenon in industry.

From the emerging work-culture literature (Kanter, 1983), we know that culture is so powerful that it can either stimulate or repress performance. That is, the strength of an organization's goals and symbol systems, the quality of group sharing and networking, the patterns of reward and recognition, and empowerment and entrepreneurial opportunities to shape the organization's future function as a context (work culture) that nurtures success. It seems that work culture is the social and psychological force that stimulates the direction and quality of work in an organization. The educational work place can be revitalized by stimulating and managing its culture.

The expansion of the knowledge base about organizational and human productivity over the past five years indicates that until educational leaders and support staff together assume responsibility for schooling, achievement patterns are likely to remain unchanged. The sociological-management challenge is one of empowering groups to address educational productivity.

Indeed, the literature on all kinds of productive organizations continues to affirm clearly and strongly that employee involvement in shaping the organization's direction is essential to the very survival of the organization. Opportunity, information, and resources are the vital materials that fuel organizational productivity.

Considering this context relative to faculty development and faculty assessment, we developed a summative–evaluation instrument for use in the college/university community. Indeed, any long–term educational operation should be in a constant state of examination for self–improvement. For this reason, colleges and universities should continually provide for their faculty evaluation. Accordingly there should be some system of feedback designed to assess and evaluate on the basis of collected data. Those responsible for making decisions must have access to data which permit them to determine faculty performance. Without such information, faculty effectiveness often cannot improve.

Recently a program evaluation committee in the College of Education at a large Southwestern State University (student enrollment of 25,000) developed a faculty–evaluation scale. Using the input of faculty and students, an assessment scale was designed. The final draft of the faculty and student survey instrument consisted of a set of fifteen teacher competencies thought to be important elements of effective teaching. The degree of emphasis was shown by a rating scale with the greater emphasis being shown by the higher rating.

Instrument Design and Methodology

The construction and field testing of an instrument involves the following steps: (a) development of a pool of instrument questions, (b) development of questionnaire forms, (c) selection of individuals to respond to the instrument, (d) actual administration of the instrument, and (e) writing computer programs for the reliability and validity assessments

of the instrument.

The questions were written using the standard domain-sampling model supported by Nunnally (1967) and Cronbach (1970). Initial questions addressing the teacher competency assessed were written, scrutinized, and studied. Next, the pool of items was expanded. Finally, the questions were rewritten and revised. These final revisions produced the pool of items contained in the instrument.

An instrument form employing a Likert format was developed for the pool of items. Lemon (1973) wrote that much of the experimental research pertaining to attitude change involves measuring attitudes by a simple, easily administered instrument, usually a rating-scale format, and then attributing changes or differences in the ratings to experimental manipulations. In this context, the Likert scale of summative ratings is the most widely used attitude measure in current research. Gronland (1976) stated that Likert's approach to attitude scaling was less time consuming in development than earlier methods such as Thurston's and Remmers'. Borg and Gall (1974) identified the Likert technique as the easiest method presently available for developing scales needed in attitude research.

McGinnies (1970) clarified the procedures involved in constructing a Likert scale. The first step involves the writing of a number of statements dealing with the topic under investigation. Respondents then indicate their reaction to each question on an interval scale. Severy (1974) added that it is assumed when a Likert scale is developed that all written questions reflect the same attitudinal dimensions. Advantages of the Likert scale are: (1) that construction is less time consuming and easier than are other scaling techniques, and (2) that fewer items are required to obtain higher correlation coefficients (Ahmann & Glock,

1975; Borg & Gall, 1974; Gronland, 1976; Lemon, 1973; Shaw & Wright, 1967).

Shaw and Wright (1967) wrote that items should be written in simple and clear language, should contain a single idea, and should be unambiguous. Borg and Gall (1974) noted that an instrument dealing with attitudes would generally be constructed as an attitude scale and should include at least ten items to obtain a valid picture of the attitude concerned. Tittle and Hill (1967) investigated the predictive efficiency of four frequently used attitude measurement techniques in terms of predicting indices of voter behavior. The Thurston, Guttman, Likert and semantic differential scales were compared. The Likert scale was superior to all other scales in predictive validity. Their evidence suggested that in cases where the Likert and Thurston scales were of equal length, the Likert scale exhibited higher reliability.

Jenkins and Taber (1977) found in an agree/disagree context that the number of response categories above five did not, in any situation, yield a significant increase in Likert discriminability. McKelvie (1978) proposed using five or six categories. He further suggested that there was no psychometric advantage in a large number of scale categories and, on the other hand, that discriminative power and validity may be reduced when fewer than five categories are used. Ramsey (1973) studied the effect of the number of categories in rating scales on the precision of scale values estimated by maximum likelihood techniques and concluded that using seven or more categories provides very nearly as much precision of estimate as a corresponding task requiring continuous judgment. Neumann and Neumann (1981) concluded from their research that the five-point scale appears to be the most convenient to use in attitudinal surveys. In addition to the fact that the literature suggestions

are frequent favoring a five- or six-point scale for the Likert format (Johnson, 1990), in our work we have found that the following six-category response choices recommended by Rotter (1972) seem to reflect relatively equidistant psychological order: 1=disagree strongly; 2=disagree; 3=tend to disagree; 4=tend to agree; 5=agree; and 6=agree strongly.

Instrument Scale Descriptors and Sampling Procedures

In constructing the response columns for the scale, the attempt was made to reflect equidistant order in the categories. In the continuum from "agree strongly" to "disagree strongly", the categories were ordered.

There is the question of a metric base for the variables. Stevens (1946) wrote that factor analysis requires the variables be measured on at least an interval level. This requirement is implied by the use of correlation or covariance matrices as the basic input to factor analysis. If it were argued that the scaling used was not absolutely of a metric base, would a researcher always avoid factor analysis? Kim and Mueller (1978) have written that such is not necessarily so. Many variables, such as measure of attitudes and opinions, do not have a clearly established metric base. Furthermore, it is generally assumed that many ordinal variables maybe given numeric values without distorting the underlying properties. The final answer to the question of applying multi-variate analysis to a nonmetric variable base really hinges on two considerations: (1) how well the arbitrarily assigned numbers reflect the underlying true distances, and (2) the amount of distortion introduced in the correlation procedures (which become the basic input to factor analysis) by the distortions of scaling. Fortunately, the correlation coefficients are fairly robust with respect to ordinal

distortions in measurement (Kim, 1975; Labovitz, 1967, 1970). Hence, as long as one can assume that the distortions introduced by assigning numeric values to ordinal categories are not very substantial, treating ordinal variables as if they were metric variables can be justified (Kim & Mueller, 1978). The scale descriptions were written to reflect as near a metric base as possible. It was felt the distortions were minimal and the application of factor analysis to the data would be consistent with factor analytic models used in instrument development.

The number of cases used in the instrument validation was 3003. The standard is to have at least five times as many subjects as try-out items (Lemke & Wiersma, 1976; Nunnally, 1970). The respondents populations were as follows:
(1) all students (n=3003); (2) educational psychology students (n=376); (3) elementary education students (n=1007); (4) educational administration students (n=208); (5) educational specializations students (n=950); and (6) secondary education students (n=422).

Final Scale Analysis of the Instrument

The collected pool of data was submitted to factor and subscale analyses (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Instrument validity was established by two methods. First content validity involves essentially the systematic examination of the scale content to determine whether it covers a representative sample of the domain to be measured (Anastasi, 1976; Ebel, 1956; Huddleston, 1956; Lennon, 1956). Representative selection of items and sensible test construction are cited by Nunnally (1967) as the standard means of ensuring content validity.

Next, construct validity was established. The construct validity of a test is the extent

to which the test may be said to measure a theoretical construct or trait (Anastasi, 1976). Essential to the notion of construct validity is the internal consistency of the items (Nunnally, 1967). Of particular relevance to construct validity is factor analysis. Lemke and Wiersma (1976) have defined factor analysis as the refined technique for analyzing the interrelationship of behavior data, simplifying the description of data by reducing the number of necessary variables or dimensions (Anastasi, 1976; Kim & Mueller, 1978).

Nunnally (1967) noted that some researchers refer to construct validity as "factorial validity." Furthermore, factor analysis is an attractive method for evaluating validity because it focuses on the reliable components of collected data (Gorsuch, 1983). Thompson (1989) noted that the "common variance" represented by indices of association tended to represent reliable variance, and since it is from these indices that factors are extracted, it follows that factors tend to be constructed from the "true score" components of variables.

Determining the number of factors to extract from the correlation matrix is a fundamental decision in any analysis (Thompson & Borrello, 1986). Many researchers follow the recommendations of Guttman (1954) and extract all factors with eigenvalues greater than one. Other researchers feel the Scree test is superior to other methods for determining the number of factors to extract (Cattell, 1965; Linn, 1968; Tucker, Koopman & Linn, 1969). These two methods have now been rigorously studied, and most of the leaders in factor analysis favor the Scree test.

We used principal components analysis (Stevens, 1986) followed by varimax rotation. Scale items were retained in the instrument only if they had a factor loading greater then or equal to 0.40. The rule often used in this context is to consider factor loading less than

0.30 as not substantial (Gorsuch, 1983; Kim and Mueller, 1978).

Evaluation of the reliability of the instrument was accomplished through use of the Kuder-Richardson generalized reliability formula, coefficient alpha (Cronbach, 1951; Ebrel, 1965; Novick & Lewis, 1967). This formula was appropriate since a Likert scaling format was employed in the instrument form.

The factor loadings for the instrument questions, as well as the Cronbach alpha reliability measures follow. Table 1 summarizes the psychometric data for the population sampled. Table 2 reports the reliability measures for each sample.

TABLE I
Summary Psychometric Data For The Scale

Area	Subjects	Instrument Questions	Factors	Factor Eigenvalue
All Respondents	3003	1-15	1	9.07
Educational Psychology	376	1-15	1	8.33
Elementary Education	1007	1-15	1	10.22
Educational Administration	208	1-4,6,9-15	1	7.48
Educational Specializations	950	1,4-15	1	8.55
Secondary Education	422	1-15	1	8.01

TABLE 2
Reliability Measures For The Scale

Area	Subjects	Instruments Questions	Cronbach's Alpha
All Respondents	3003	1-15	0.95731
Educational Psychology	376	1-15	0.94807
Elementary Education	1007	1-15	0.96920
Educational Administration	208	1-4,6,9-15	0.92488
Educational Specializations	950	1,4-15	0.95161
Secondary Education	422	1-15	0.94296

For the educational administration area, the data suggest that the scale is comprised of questions 1 through 4, 6 and 9 through 15. For the educational specializations area, the data suggest that the scale consist of questions 1 and 4 through 15.

In addition, the appendices to this document report the following information for each population surveyed: (1) descriptive statistics for the instrument variables (questions); (2) the correlation matrix for the instrument variables; and (3) the factor matrix using principal factoring (PA2) with interactions. See appendices A through R for this information. Finally the instrument variables are reported in Appendix S.

Summary and Conclusion

This article illustrated the process for the development of an instrument for faculty evaluation. The process linked the literature review with a psychometric analysis of the collected data.

The instrument was shown to be reliable and valid. Because of the reliability measures, the instrument can be used either for individual or group assessment. If it were desired that norms be established for each of the populations studied, such would involve the use of converting and equating procedures to generate a normal distribution with a specified mean and standard deviation. This process would allow the generation of percentile reference points for the interpretation of individual scores. A sample could be checked with respect to student classification, socioeconomic level, or other information relevant to the study. For this particular study, however, norms were not developed. Such could be a point-of-departure for another study which would encompass a separate body of psychometric literature and statistical analyses.

References

Ahmann, J.S., & Glock, M.D. (1975). **Evaluating pupil growth.** Boston: Allyn & Bacon.

Anastasi, A. (1976). **Psychological testing** (4th ed.) New York: Macmilliam.

Borg, W.R., & Gall, M. (1974). **Educational research.** New York: David McKay Company.

Brookover, W.B., & Lezotte, L.W. (1979). **Changes in school characteristics coincident with changes in student achievement.** East Lansing: Michigan State University, College of Education. (ERIC Document Reproduction Service No. ED 181 005).

Cattell, R.B. (1965). Factor analysis: an introduction to essentials (I) the purpose and underlying models, (II) the role of factor analysis in research. **Biometrics,** 21, 190-215, 405-435.

Coulson, J.E. (1977, July). **Overview of the national evaluation of the emergency school and aid act.** Santa Monica, CA: System Development Corporation (ED 154 951)

Cronbach, J.E. (1951). Coefficient alpha and the internal structure of tests. **Psychometrika,** 16, 3, 297-334.

Cronbach, L.J. (1970). **Essentials of psychological testing.** New York: MacMillan.

Ebel, R.L. (1956). Obtaining and reporting evidence on content validity. **Educational and Psychological Measurement,** 16, 3, 169-282.

Ebel, R.L. (1965). **Measuring educational achievement.** Englewood Cliffs, NJ: Prentice-Hall.

Edmonds, R.R. (1979). Effective schools for the urban poor. **Educational Leadership** 37, 15-23.

Edmonds, R.R. (December 1982). Programs for school improvement: An overview. **Educational Leadership,** 40, 4-11.

Glenn, B.C. (1981). **What works? An examination of effective schools for poor black children.** Cambridge: Harvard University, Center for Law and Education

Gorsuch, R.L. (1983). **Factor analysis.** Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gronland, N.E. (1976). **Measurement and evaluation in teaching.** New York: MacMillan.

Guttman, L. (1954). Some necessary conditions for common factor analysis. **Psychometrika,** 19, 149-161.

Huddelston, E.M. (1956). Test development on the basis of content validity. **Educational and Psychological Measurement,** 16, 3, 283-293.

Jenkins, G.D., & Taber, T.D. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. **Journal of Applied Psychology,** 62, 392-398.

Johnson, W.L. (1990). Factorial structure of the Charles F. Kettering Climate Profile. **Psychological Reports,** 67, 67-74.

Johnson, W.L., & Snyder, K.J. (1986). Instructional leadership effectiveness: A research analysis and strategy. **Educational and Psychological Research,** 6,1, 27-47.

Kanter, R.M. (1983). **The change masters: Innovation for productivity in the American corporation.** New York: Simon & Schuster

Kilmann, R., Saxton, M.J. & Serpa, R. (1985). **Gaining control of the corporate structure.** San Francisco: Jossey-Bass.

Kim, J.O. (1975). Mulitvariate analysis of ordinal variables. **American Journal of Sociology,** 81, 261-298.

Kim, J.O., & Mueller, C.W. (1978). **Factor Analysis.** Beverly Hills, CA.: Sage Publications.

Klausemeier, H.J. (1982, September). "A research strategy for educational improvement." **Educational Digest,** 48, 1, 19-22.

Labovitz, S. (1967). Some observations on measurement and statistics. **Social Forces,** 46, 151-160.

Labovitz, S. (1970). The assignments of numbers to rank order categories. **American Sociological Review,** 35, 515-524.

Lemke, E., & Wiersma, W. (1976). **Principles of psychological measurement.** Chicago, IL: Houghton-Mifflin.

Lemon, N. (1973). **Attitudes and their measurement.** London: B.T. Batsford.

Lennon, R.T. (1956). Assumptions underlying the use of content validity. **Educational and Psychological Measurement,** 16, 3, 294-304.

Lieberman, A., & Miller, L. (1978). Introduction to **Staff development: New demands, new realities, new perspectives.** New York: Teachers College Press.

Linn, R.L. (1968). A Monte Carlo approach to the number of factors problem.

Psychometrika, 33, 37-71.

Marcus, A.C., MacQueen, A.H., Duck, G.A., Lee, R.D. & Wellisch, J.B. (1976, April 14-16). **Administrative leadership in a sample of successful schools from the National Evaluation of the Emergency School.** System Development Corporation. Paper presented at the annual meeting of the AERA Convention, San Francisco, CA.

Miller, L., & Wolf, T. (1978). Staff development for school change: Theory and practice. In A. Lieberman & L. Miller (eds.), **Staff development: New demands, new realities, new perspectives** (pp. 144-160). New York: Teachers College Press.

McGinnies, E. (1970). **Social Behavior: A functional analysis.** Boston: Houghton Mifflin Company.

McKelvie, S.G. (1978). Graphic rating scales: how many categories. **British Journal of Psychology,** 69, 185-202.

Neumann L., & Neumann, Y. (1981). Comparison of six lengths of rating scales: students' attitudes toward instruction. **Psychological Reports,** 48, 399-404.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Bent, D.H. (1975). **Statistical package for the social sciences.** New York, NY: McGraw-Hill.

Novick, M.R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. **Psychometrika,** 32, 1, 1-13.

Nunnally, J.C. (1967). **Psychometric theory.** New York, NY: McGraw-Hill.

Nunnally, J.C. (1970). **Introduction to Psychological Measurement.** New York, NY: McGraw-Hill.

Peters, T.J., & Waterman, R.H. (1982). **In Search of Excellence: Lessons from America's Best Run Companies.** New York: Harper & Row.

Ramsey, J.O. (1973). The effect of number of categories in rating scales on precision of estimation of scale values. **Psychometrika,** 38, 513-532.

Rotter, G.S. (1972). Attitude points of agreement and disagreement. **Journal of Social Psychology,** 86, 211-218.

Severy, L.J. (1974). **Procedure and issues in the measurement of attitudes.** Washington, D.C.: National Institute of Education. (ERIC Document Reproduction Service No. ED 099 426).

Shaw, M.E., & Wright, J.M. (1967). **Scales for the measurement of attitudes.** New York: McGraw-Hill.

Snyder, K.J. & Anderson, R.H. (1986). **Managing productive schools: Toward an ecology.** San Diego: Harcourt Brace & Jovanovich, Inc.

Snyder, K.J. (1988). **Managing productive schools.** San Diego: Harcourt Brace & Jovanovich, Inc.

Stevens, J. (1986). **Applied Multivariate Statistics for the Social Sciences.** Hillsdale, N.J.: Lawrence Erlbaum & Associates, Inc.

Stevens, S.S. (1946). On the theory of scales measurement. **Science,** 103, 677-680.

Thompson, B. (1989). Meta-analysis of factor structures: a case study example with Bem's androgyny measure. **Journal of Experimental Education,** 57, 187-197.

Thompson, B., & Borrello, G.M. (1986). Second-order factor structure of the MBTI: a construct validity assessment. **Measurement and Evaluation in Counseling and Development,** 18, 148-153.

Tittle, C.R., & Hill, R.E.J. (1967). Attitude measurement and prediction of behavior: An evaluation of conditions and measurement techniques. **Sociometry,** 30, 2, 199-213.

Tucker, L.R., Koopman, R.F., & Linn, R.L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. **Psychometrika,** 34, 421-459.

APPENDIX A
Descriptive Statistics for the Instrument
Variables (Questions)
All Respondents (n=3003)

Variables	M	Mdn	Mo	SE	SD
Q1	3.419	3.610	4.000	0.014	0.780
Q2	3.506	3.617	4.000	0.012	0.661
Q3	3.502	3.648	4.000	0.013	0.695
Q4	3.416	3.598	4.000	0.014	0.794
Q5	3.526	3.687	4.000	0.013	0.699
Q6	3.538	3.718	4.000	0.013	0.725
Q7	3.570	3.718	4.000	0.012	0.662
Q8	3.717	3.847	4.000	0.011	0.589
Q9	3.483	3.648	4.000	0.013	0.730
Q10	3.475	3.636	4.000	0.013	0.726
Q11	3.375	3.541	4.000	0.014	0.787
Q12	3.360	3.526	4.000	0.014	0.793
Q13	3.405	3.619	4.000	0.015	0.816
Q14	3.505	3.690	4.000	0.014	0.741
Q15	3.444	3.690	4.000	0.016	0.851

APPENDIX B

Correlation Matrix for the Instrument Variables (Questions)
All Respondents (n=3003)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000	—													
Q2	0.4740	1.0000													
Q3	0.5256	0.4849	1.0000												
Q4	0.6230	0.4698	0.5438	1.0000											
Q5	0.6301	0.5195	0.6345	0.6391	1.0000										
Q6	0.6274	0.4950	0.6288	0.5878	0.7291	1.0000									
Q7	0.5713	0.4834	0.6825	0.5441	0.6908	0.6872	1.0000								
Q8	0.5618	0.4857	0.5450	0.5229	0.6184	0.6307	0.6059	1.0000							
Q9	0.6102	0.4658	0.5391	0.5818	0.6171	0.6324	0.5819	0.6587	1.00000						
Q10	0.6314	0.4850	0.5771	0.5758	0.6367	0.7100	0.6154	0.6367	0.66069	1.00000					
Q11	0.6042	0.5149	0.5324	0.6411	0.6317	0.6064	0.5872	0.5728	0.59338	0.62355	1.0000				
Q12	0.5936	0.4371	0.5355	0.6155	0.6166	0.5973	0.5737	0.5302	0.56577	0.58610	0.58784	1.0000			
Q13	0.6321	0.4542	0.5356	0.5996	0.6445	0.6706	0.5777	0.6258	0.68260	0.66729	0.60672	0.65482	1.0000		
Q14	0.6072	0.4470	0.5361	0.5809	0.6511	0.6748	0.5966	0.6324	0.69070	0.66747	0.58197	0.61715	0.81044	1.00000	
Q15	0.6642	0.4650	0.5689	0.6130	0.6758	0.7099	0.6191	0.6373	0.66280	0.69451	0.62555	0.62829	0.77446	0.75430	1.0000

APPENDIX C

Factor Matrix Using Principal Factoring (PA2) with Iterations
All Respondents (n=3003)

Variables	Factor 1*
Q1	0.769
Q2	0.604
Q3	0.719
Q4	0.746
Q5	0.825
Q6	0.833
Q7	0.774
Q8	0.760
Q9	0.788
Q10	0.810
Q11	0.762
Q12	0.748
Q13	0.830
Q14	0.821
Q15	0.845

*Eigenvalue of 9.07

APPENDIX D

Means and Standard Deviations for the Instrument Variables (Questions)
Educational Psychology Area (n=376)

Variables	M	SD
Q1	3.5426	0.6915
Q2	2.5319	0.6274
Q3	3.5293	0.6999
Q4	3.4681	0.7859
Q5	3.5904	0.6510
Q6	3.6223	0.6661
Q7	3.6250	0.6494
Q8	3.7846	0.5048
Q9	3.5106	0.6927
Q10	3.5080	0.6850
Q11	3.4255	0.7191
Q12	3.4840	0.7301
Q13	3.5133	0.7999
Q14	3.5559	0.7133
Q15	3.5372	0.7720

APPENDIX E

Correlation Matrix for the Instrument Variables (Questions) for the Educational Foundations Area (n=376)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.4947	1.0000													
Q3	0.4685	0.4077	1.0000												
Q4	0.5913	0.4239	0.4695	1.0000											
Q5	0.5897	0.5021	0.5999	0.5842	1.0000										
Q6	0.5155	0.4692	0.5386	0.5118	0.6816	1.0000									
Q7	0.4780	0.4385	0.6843	0.4389	0.6387	0.5718	1.0000								
Q8	0.4961	0.4385	0.4670	0.4363	0.5747	0.6060	0.4688	1.0000							
Q9	0.5835	0.3980	0.4476	0.5835	0.5063	0.5346	0.4920	0.5594	1.0000						
Q10	0.5199	0.4492	0.5001	0.4635	0.5395	0.6319	0.4893	0.5718	0.5420	1.0000					
Q11	0.5480	0.4781	0.4255	0.5572	0.5157	0.4922	0.5139	0.4515	0.4994	0.5127	1.0000				
Q12	0.5823	0.5075	0.5253	0.5335	0.6650	0.5962	0.5807	0.5079	0.5433	0.5360	0.5615	1.0000			
Q13	0.5895	0.4375	0.5185	0.5712	0.5891	0.6501	0.5358	0.6246	0.5892	0.6178	0.5001	0.6738	1.0000		
Q14	0.5708	0.3862	0.5308	0.5097	0.5891	0.6281	0.5778	0.5778	0.5788	0.6102	0.5461	0.6493	0.7791	1.0000	
Q15	0.6214	0.4655	0.5877	0.5865	0.6406	0.7016	0.5997	0.6125	0.5976	0.6625	0.5526	0.6586	0.8218	0.7879	1.0000

APPENDIX F

Factor Matrix using Prinicpal Factoring (PA2) with Iterations for the Educational Psychology Area (n=376)

Variables	Factor 1*
Q1	0.733
Q2	0.595
Q3	0.686
Q4	0.694
Q5	0.792
Q6	0.786
Q7	0.719
Q8	0.710
Q9	0.714
Q10	0.735
Q11	0.681
Q12	0.784
Q13	0.829
Q14	0.811
Q15	0.870

*Eigenvalue of 8.23

APPENDIX G

Means and Standard Deviations for the Instrument Variables (Questions)
Elementary Education Area (n=1007)

Variables	M	SD
Q1	3.3118	0.8643
Q2	3.4300	0.7372
Q3	3.4012	0.7748
Q4	3.3128	0.8509
Q5	3.4320	0.7920
Q6	3.4141	0.8247
Q7	3.4677	0.7531
Q8	3.6058	0.7053
Q9	3.3774	0.8091
Q10	3.4022	0.7989
Q11	3.2721	0.8745
Q12	3.2175	0.8503
Q13	3.2890	0.8833
Q14	3.4320	0.7995
Q15	3.3178	0.9386

APPENDIX H

Correlation Matrix for the Instrument Variables (Questions)
Elementary Education Area (n=1007)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.5964	1.0000													
Q3	0.6032	0.5834	1.0000												
Q4	0.7383	0.6029	0.6356	1.0000											
Q5	0.7214	0.6535	0.7005	0.7063	1.0000										
Q6	0.7327	0.6288	0.6777	0.5607	0.7849	1.0000									
Q7	0.6589	0.6166	0.7256	0.6400	0.7523	0.7632	1.0000								
Q8	0.6622	0.6378	0.6371	0.6279	0.7268	0.7354	0.7122	1.000							
Q9	0.6822	0.5908	0.6144	0.6599	0.7039	0.7144	0.6642	0.7590	1.0000						
Q10	0.7315	0.6174	0.6672	0.6920	0.7415	0.7818	0.7310	0.7491	0.7291	1.0000					
Q11	0.6891	0.6246	0.6411	0.7203	0.7213	0.7008	0.6939	0.6558	0.6666	0.7167	1.0000				
Q12	0.6670	0.5340	0.5795	0.6820	0.6573	0.6141	0.6046	0.5856	0.6087	0.6671	0.6408	1.0000			
Q13	0.6910	0.5828	0.5915	0.6810	0.7051	0.7211	0.6467	0.6840	0.7290	0.7211	0.6804	0.6692	1.0000		
Q14	0.6772	0.5918	0.6040	0.6675	0.7096	0.7218	0.6627	0.7270	0.7419	0.7252	0.6349	0.6307	0.8350	1.0000	
Q15	0.7398	0.5722	0.6268	0.7005	0.7324	0.7607	0.6823	0.7224	0.6900	0.7400	0.7191	0.6468	0.8075	0.7613	1.0000

APPENDIX I

Factor matrix using Principal Factoring (PA2) with Iterations for the Elementary Education Area (n=1007)

Variables	Factor 1*
Q1	0.83142
Q2	0.72402
Q3	0.76571
Q4	0.81388
Q5	0.87073
Q6	0.87025
Q7	0.82757
Q8	0.83465
Q9	0.82848
Q10	0.87085
Q11	0.82267
Q12	0.75682
Q13	0.84672
Q14	0.84167
Q15	0.86164

*Eigenvalue of 10.22

APPENDIX J

Means and Standard Deviations for the Instrument Variables (Questions) Educational Administration Area (n=208)

Varibles	M	SD
Q1	3.5865	0.6311
Q2	3.6635	0.5665
Q3	3.5913	0.6455
Q4	3.4567	0.8094
Q5	3.6058	0.6046
Q6	3.6346	0.6454
Q7	3.6538	0.5777
Q8	3.8029	0.3988
Q9	3.6106	0.6647
Q10	3.5481	0.6502
Q11	3.4038	0.7226
Q12	3.4615	0.7540
Q13	3.5481	0.6934
Q14	3.6058	0.6281
Q15	3.6587	0.6622

APPENDIX K

Correlation Matrix for the Instrument Variables (Questions)
Educational Administration Area (n=208)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3655	1.0000													
Q3	0.4488	0.4411	1.0000												
Q4	0.2958	0.3262	0.4144	1.0000											
Q5	0.2924	0.1890	0.5135	0.4091	1.0000										
Q6	0.4575	0.3887	0.7297	0.4597	0.5451	1.0000									
Q7	0.3210	0.2032	0.6551	0.2467	0.5755	0.5790	1.0000								
Q8	0.4424	0.2181	0.4549	0.3102	0.5376	0.5259	0.5621	1.0004							
Q9	0.4089	0.3430	0.4717	0.4579	0.5056	0.5563	0.3266	0.4744	1.0000						
Q10	0.5078	0.4375	0.6397	0.4491	0.4662	0.6866	0.4946	0.5118	0.5297	1.0000					
Q11	0.4526	0.2982	0.3969	0.5008	0.4988	0.4422	0.3943	0.4787	0.5503	0.5137	1.0000				
Q12	0.3724	0.3201	0.5878	0.5157	0.4963	0.6459	0.4572	0.4325	0.5241	0.6146	0.4808	1.0000			
Q13	0.4099	0.3734	0.6323	0.4556	0.4717	0.6871	0.4880	0.5323	0.6644	0.6806	0.5010	0.6595	1.0000		
Q14	0.5008	0.3720	0.5776	0.4318	0.4410	0.7154	0.4075	0.4982	0.6718	0.6616	0.4589	0.6308	0.7868	1.0000	
Q15	0.5507	0.4134	0.5648	0.3373	0.3983	0.5884	0.3589	0.3659	0.5416	0.6161	0.4409	0.5202	0.6093	0.6389	1.0000

APPENDIX L

Factor Matrix Using Principal Factoring (PA2) with Iterations for the Educational Administration Area (n=208)*

Variables	Factor 1
Q1	0.587
Q2	0.472
Q3	0.777
Q4	0.565
Q5	0.651
Q6	0.834
Q7	0.658
Q8	0.649
Q9	0.723
Q10	0.807
Q11	0.642
Q12	0.749
Q13	0.832
Q14	0.818
Q15	0.717

*Eigenvalue for Factor 1 is 7.48334.

APPENDIX M

Means and Standard Deviations for the Instrument Variables (Questions)
Educational Specializations Area (n=950)

Variables	M	SD
Q1	3.4684	0.7462
Q2	3.5642	0.5946
Q3	3.5505	0.6047
Q4	3.4937	0.7254
Q5	3.5568	0.6512
Q6	3.5684	0.6739
Q7	3.6116	0.5894
Q8	3.7747	0.5232
Q9	3.5589	0.6779
Q10	3.4905	0.7166
Q11	3.4421	0.7246
Q12	3.4189	0.7649
Q13	3.4674	0.7588
Q14	3.5453	0.7269
Q15	3.4705	0.8343

APPENDIX N

Correlation Matrix for the Instrument Variables (Questions)
Educational Specializations Area (n=950)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3441	1.0000													
Q3	0.4950	0.4158	1.0000												
Q4	0.6370	0.3478	0.5087	1.0000											
Q5	0.6627	0.4314	0.5961	0.6487	1.0000										
Q6	0.5993	0.3558	0.5680	0.5785	0.7018	1.0000									
Q7	0.5602	0.3944	0.6390	0.5426	0.6547	0.6386	1.0000								
Q8	0.5053	0.3242	0.4323	0.4682	0.5077	0.4830	0.5155	1.0000							
Q9	0.6149	0.3042	0.4849	0.5353	0.5999	0.5723	0.5623	0.5989	1.0000						
Q10	0.6279	0.3265	0.4825	0.5613	0.6063	0.6787	0.5513	0.5395	0.6583	1.0000					
Q11	0.5909	0.4378	0.5020	0.6207	0.5987	0.5616	0.5481	0.5242	0.5518	0.6005	1.0000				
Q12	0.6083	0.3114	0.4849	0.6257	0.6058	0.5964	0.5997	0.4809	0.5375	0.5493	0.5913	1.0000			
Q13	0.6698	0.3047	0.4467	0.5681	0.6584	0.6152	0.5312	0.5653	0.6612	0.6688	0.5839	0.6226	1.0000		
Q14	0.6183	0.2943	0.4574	0.5660	0.6846	0.6507	0.5858	0.5698	0.6659	0.6632	0.5600	0.6101	0.8058	1.0000	
Q15	0.6711	0.3224	0.4968	0.5976	0.6634	0.6650	0.5884	0.5858	0.6652	0.7096	0.5879	0.6186	0.7540	0.7630	1.0000

APPENDIX O

Factor Matrix Using Principal Factoring (PA2) With Iterations* for the Educational Specializations Area (n=950)*

Variables	Factor 1
Q1	0.78135
Q2	0.45678
Q3	0.66740
Q4	0.74328
Q5	0.82239
Q6	0.79184
Q7	0.75526
Q8	0.67143
Q9	0.76848
Q10	0.79062
Q11	0.74429
Q12	0.74838
Q13	0.82749
Q14	0.82983
Q15	0.84443

*Eigenvalue for Factor 1 is 8.55774.

APPENDIX P

Means and Standard Deviations for the Instrument
Variables (Questions)
Secondary Education Area (n=422)

Variables	M	SD
Q1	3.3720	0.7591
Q2	3.4431	0.6649
Q3	3.5592	0.6858
Q4	3.4597	0.7209
Q5	3.5853	0.6362
Q6	3.6209	0.6453
Q7	3.6232	0.6069
Q8	3.7536	0.5435
Q9	3.4763	0.6948
Q10	3.5308	0.6300
Q11	3.4384	0.7485
Q12	3.3815	0.7606
Q13	3.3697	0.8276
Q14	3.4858	0.7053
Q15	3.4882	0.7970

APPENDIX Q

Correlation Matrix for the Instrument Variables (Questions)
Secondary Education Area (n=422)

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Q1	1.0000														
Q2	0.3549	1.0000													
Q3	0.4069	0.3980	1.0000												
Q4	0.4636	0.4758	0.5164	1.0000											
Q5	0.5070	0.4241	0.5762	0.5978	1.0000										
Q6	0.4874	0.4090	0.5982	0.5593	0.6979	1.0000									
Q7	0.4544	0.3499	0.6215	0.5487	0.6462	0.6532	1.0000								
Q8	0.4069	0.4014	0.5299	0.4534	0.5005	0.5930	0.5314	1.0000							
Q9	0.4603	0.4777	0.5013	0.5291	0.5015	0.5996	0.5279	0.5694	1.0000						
Q10	0.4602	0.4634	0.5262	0.4812	0.4971	0.6130	0.5366	0.6771	0.6093	1.0000					
Q11	0.5023	0.4725	0.4420	0.5719	0.5821	0.5956	0.4898	0.5230	0.5566	0.5631	1.0000				
Q12	0.4200	0.3787	0.4642	0.5846	0.5043	0.4840	0.4407	0.4692	0.5137	0.4240	0.5065	1.0000			
Q13	0.4687	0.3836	0.4928	0.5585	0.5309	0.6322	0.5143	0.5673	0.6637	0.5430	0.5928	0.6357	1.0000		
Q14	0.4291	0.3908	0.4829	0.5407	0.5293	0.6143	0.5007	0.5405	0.6900	0.5355	0.5674	0.5436	0.7863	1.0000	
Q15	0.4607	0.4469	0.5075	0.5758	0.6296	0.6793	0.5824	0.5251	0.6557	0.5423	0.5681	0.5853	0.7304	0.7222	1.0000

APPENDIX R

Factor Matrix Using Prinicpal Factoring (PA2) With Iterations for the Secondary Education Area (n=422)*

Variables	Factor 1
Q1	0.609
Q2	0.561
Q3	0.690
Q4	0.729
Q5	0.759
Q6	0.815
Q7	0.727
Q8	0.705
Q9	0.775
Q10	0.721
Q11	0.738
Q12	0.680
Q13	0.805
Q14	0.783
Q15	0.816

*Eigenvalue for Factor 1 is 8.01242.

APPENDIX S

Questions on the Instrument

1. Has made his/her expectations clear.
2. Has maintained adequate office hours for consultation.
3. Has been tolerant of differing viewpoints regarding the subject matter.
4. Has utilized appropriate grading practices.
5. Has been responsive to student needs.
6. Has related well to students.
7. Has been open to student input (suggestions, comments, recommendations, etc.).
8. Has appeared knowledgeable in his/her field.
9. Has presented material relevant to the course description in the catalog.
10. Has communicated (speech, mannerisms, delivery, etc.) effectively.
11. Has provided appropriate feedback on student work.
12. Has given realistic class assignments.
13. Has made the course worthwhile (I have learned enough to make it worth my time and money).
14. Has presented the course in such a manner that I see present or future applications of the materials, ideas, skills, etc.
15. Has been of such caliber that I would recommend him/her to other students.